Capturing rogue waves by multi-point statistics

A. Hadjihosseini[1]**, Matthias Wächter**[1]**, N. P. Hoffmann**[2,3] **and J. Peinke**[1,4]

[1] Universität Oldenburg, 26111 Oldenburg, Germany
[2] Hamburg University of Technology, 21073 Hamburg, Germany
[3] Imperial College, London SW7 2AZ, United Kingdom
[4] Fraunhofer Institute for Wind Energy and Energy System Technology,
Ammerlander Heerstr. 136, DE-26129 Oldenburg, Germany

E-mail: ali.hadjihosseini@uni-oldenburg.de

Abstract. As an example for complex systems with extreme events we investigate ocean wave states exhibiting rogue waves. We present a statistical method of data analysis based on multi-point statistics which for the first time allows grasping extreme rogue wave events in a statistically highly satisfactory manner. The key to the success of the approach is mapping the complexity of multi-point data onto the statistics of hierarchically ordered height increments for different time scales for which we can show that a stochastic cascade process with Markov properties is governed by a Fokker-Planck equation. Conditional probabilities as well as the Fokker-Planck equation itself can be estimated directly from the available observational data. With this stochastic description surrogate data sets can in turn be generated allowing to work out arbitrary statistical features of the complex sea state in general and extreme rogue wave events in particular. The results also open up new perspectives for forecasting the occurrence probability of extreme rogue wave events, and even for forecasting the occurrence of individual rogue waves based on precursory dynamics.

1. Introduction and motivation

The observation and study of waves on the sea is probably one of the oldest scientific and cultural endeavours of mankind. But even today the sea's state can not be regarded as anything else than an enigma to man and science. Of course ocean waves have inspired a tremendous number of often groundbreaking results in mathematics, physics and related sciences, including nonlinear waves, localisation, extreme events, turbulence and many more. But still, the fully irregular and complex state of the sea is far from being understood. And both the characteristics of its irregularity, as well as the rare but extremely large wave events occurring sometimes, now often called rogue waves, abscond satisfying description, even in statistical terms.

Obviously the difficulties with understanding irregular and extreme or rogue ocean waves has to be seen in the context of extreme events in complex systems in general. Driven by various motives there has been extensive research on extreme events in a many fields, from the sciences, via meteorology and climate change, up to the social sciences and economics [1, 2, 3, 4, 5, 6]. It is still a strongly debated question whether extreme events are generally linked to some universal stochastic mechanisms or if they rather originate through special features of the individual systems under study [7]. Still, a common point of all observations is that the empirical data are frequently punctuated by extreme events which seem to play an important role. Often an analysis approach is to approximate the observations by means of a generalized stochastic model in which some variables are represented in terms of stochastic components [8]. Usually the complex systems under study are very high dimensional and thus finding adequate methods to model the stochastic components remains a challenge.

Besides the description of extreme events in complex systems there is also the demand of their prediction. Despite the fact that we have irregular and complex behavior of rogue waves, there are increasing number of research towards defining an early warning system for rogue wave occurrences [9, 10] or establishing a prediction method for short term prediction of rogue events. Studies on prediction methods mainly relys on deterministic behavior of non-stationary solutions of the underlying wave equations [11, 12, 13] and also deterministic nonlinear time series analysis [14].

The present work is based on the finding that complex systems can often be described highly successfully as stochastic processes in scale rather than time or space. Examples are now known for various very different fields, like turbulence[15, 16], economics[17, 18, 19], biology[20, 21, 22], and many more, see [23]. With scale dependent processes, originally introduced by Friedrich and Peinke [15], fractal and multi-fractal structures [23] and even more generally joint n-point statistics [18, 24] can be reduced by Markov properties to particular three point statistics. In a previous study [25] we have already shown for ocean wave data that certain scale dependent processes may have Markov properties. However, in [25] the Markov properties for the pure scale process

could only be derived for deliberately pre-filtered data. In the present contribution we extend our previous investigations and base our analysis of the wave dynamics on general joint $(N+1)$-point probability density functions (PDF) $p(h(t), h(t - \tau_1), ..., h(t - \tau_N))$. Here $h(t)$ denotes the water surface elevation measured at a given location at time t and τ_i are different time increments. The joint PDF provides the likelihood of a sequence of water surface elevation heights for $N + 1$ different instants of time. We show how a Fokker-Planck equation can be derived which describes these general joint PDFs. Knowing the corresponding stochastic process for the general multi-point statistics we can show that also the extreme events, i.e. the rogue waves, are grasped by this stochastic approach. The approach also allows time-series reconstruction in a statistical sense, and thus a statistically valid prediction of rogue wave occurrence.

The paper is structured as follows. First the mathematical aspects of multi-point and multi-scale description as well as the connection to scale dependent stochastic processes are introduced. Then the validity of the description based on observational ocean wave data is demonstrated. Finally the approach is applied to reconstruct time series for the underlying observational data and to forecast the occurrence probability of rogue waves in the given sea state.

2. Multi-point statistics

In this section the statistical background of our approach for a multi-point reconstruction is presented. In the following we use the short-hand notation $h_i := h(t_i)$ for the elevation of the water surface measured at a given location at time t_i, with $h_{i+j} := h(t_i + t_j)$. We define the relative change in surface height over a time interval or, respectively, a time scale τ_j as

$$\xi_j \equiv \xi(\tau_j) := h(t_i) - h(t_i - \tau_j). \tag{1}$$

The aim is to calculate the joint probability $p(h^*, t^*; h_1, t^* - \tau_1; ...; h_N, t^* - \tau_N)$ of occurrence of the event $\{h^*, t^*\}$, together with the knowledge of the past points $\{h_1, t^* - \tau_1; h_2, t^* - \tau_2; ...; h_N, t^* - \tau_N\}$. We assume that the system has no explicit time dependence, i.e. the system is stationary. The probability of occurrence of the event $\{h^*\}$ under the conditions of the past points is given by

$$p(h^* | h_1, \tau_1; ...; h_N, \tau_N) = \frac{p(h^*; h_1, \tau_1; ...; h_N, \tau_N)}{p(h_1, \tau_1; ...; h_N, \tau_N)}. \tag{2}$$

Next, the joint (N+1)-point PDF can be expressed in an equivalent way by joint increments statistics

$$p(h^*; h_1, \tau_1; h_2, \tau_2; ...; h_N, \tau_N) = p(h^* - h_1, \tau_1; h^* - h_2, \tau_2; ...; h^* - h_N, \tau_N, h^*)$$
$$= p(\xi_1; \xi_2; ...; \xi_N, h^*)$$
$$= p(\xi_1; \xi_2; ...; \xi_N | h^*) \cdot p(h^*). \tag{3}$$

Note instead of the knowledge of wave heights at N+1 points, we consider now the knowledge of N height increments and one selected height h^*. Without loss of generality we take $\tau_i < \tau_{i+1}$, and thus introduce a hierarchical ordering of the increments ξ_i.

Only if the conditional PDFs do not depend on h^*, i. e. if

$$p(\xi_1; \xi_2; ...; \xi_N | h^*) = p(\xi_1; \xi_2; ...; \xi_N), \tag{4}$$

the (N+1)-point statistics reduces to N-scale statistics of the increments ξ_i at scales τ_i. In our previous work [25] we had applied filtering based on Hilbert-Huang transform techniques (HHT) to the wave data. The filtering removed the dependency on h^* by kind of separating off the underlying dominant frequency, i.e. the wave like nature of the system. Still, already in this case Markov properties could be shown for the filtered wave data and thus the multi-scale PDF could be factorized in

$$p(h^*; h_1, \tau_1; ...; h_N, \tau_N) = p(\xi_1; ...; \xi_N) \cdot p(h^*)$$
$$= p(\xi_1 | \xi_2) \cdot ... \cdot p(\xi_{N-1} | \xi_N) \cdot p(\xi_N) \cdot p(h^*). \tag{5}$$

In our present work here we do not apply any pre-filtering and stay focussed on the very direct data itself. This renders the approach much more general and we directly start with the multi-point statistics (Eq. (2)) to investigate if the Markov property of the process is given for

$$p(\xi_j | \xi_{j+1}, \xi_{j+2}, ..., \xi_{j+N}, h^*) = p(\xi_j | \xi_{j+1}, h^*). \tag{6}$$

More specifically we will first investigate from the observational data if

$$p(\xi_j | \xi_{j+1}, \xi_{j+2}, h^*) = p(\xi_j | \xi_{j+1}, h^*) \tag{7}$$

holds. This we will take as hint for the validity of Markov property. Using Eq. (6), the multi-point PDF Eq. (3) can then be factorised as

$$p(h^*; h_1, \tau_1; h_2, \tau_2; ...; h_N, \tau_N) = p(\xi_1 | \xi_2, h^*) \cdot ... \cdot p(\xi_{N-1} | \xi_N, h^*) \cdot p(\xi_N | h^*) \cdot p(h^*). \tag{8}$$

As Eq. (6) is nothing else than the Markov property of a stochastic process of ξ_i evolving in the time scale τ_i, the evolution of conditional PDFs of Eq. (8) can be expressed by Kramers-Moyal expansion [26],

$$- \tau_j \frac{\partial}{\partial \tau_j} p(\xi_j | \xi_k, h^*) = \sum_{n=1}^{\infty} (-\frac{\partial}{\partial \xi_j})^n \left[D^{(n)}(\xi_j, \tau_j, h^*) p(\xi_j | \xi_k, h^*) \right], \tag{9}$$

where Kramers-Moyal coefficients $D^{(n)}$ are defined as

$$D^{(n)}(\xi_j, \tau_j, h^*) = \lim_{\delta_\tau \to 0} \frac{\tau_j}{n! \delta \tau} < \left[\xi_j'(\tau_j - \delta\tau, h^*) - \xi_j(\tau_j, h^*) \right]^n >_{\xi_j'}. \tag{10}$$

Note that the pre-factor $-\tau$ in Eq. (9) indicates that we consider the process for decreasing τ- values and an evolution in log-scale of τ.

If the Kramers-Moyal coefficient $D^{(4)}$ is zero, then it follows from the Pawula theorem that all coefficients for $n \geq 3$ are zero, too cf. [26]. The Kramers-Moyal expansion then yields a Fokker-Planck equation with just two coefficients,

$$-\tau_j \frac{\partial}{\partial \tau_j} p(\xi_j | \xi_k, h^*) = -\frac{\partial}{\partial \xi_j} \left[D^{(1)}(\xi_j, \tau_j, h^*) p(\xi_j | \xi_k, h^*) \right] + \frac{\partial^2}{\partial \xi_j^2} \left[D^{(2)}(\xi_j, \tau_j, h^*) p(\xi_j | \xi_k, h^*) \right]. \tag{11}$$

$D^{(1)}$ denotes the drift and $D^{(2)}$ the diffusion coefficient. With this the Fokker-Planck equation turns out a suitable description for the conditional probabilities of the water surface height increments, from which in turn the general multi-point joint PDF of the surface heights themselves, Eq. (8), can be determined as

$$p(h^*|h_1, \tau_1; ...; h_N, \tau_N) = \frac{p(h^*)}{p(h_1, \tau_1)} \times p(h^* - h_1, \tau_1 | h^* - h_2, \tau_2; h^*) \times$$

$$\times ... \times \frac{p(h^* - h_{N-1}, \tau_{N-1} | h^* - h_N, \tau_N; h^*)}{p(h_1 - h_{N-1}, \tau_{N-1} - \tau_1 | h_1 - h_N, \tau_N - \tau_1; h_1, \tau_1)}$$

$$\times \frac{p(h^* - h_N, \tau_N | h^*)}{p(h_1 - h_N, \tau_N - \tau_1 | h_1, \tau_1)}.$$

Using the increment notation, omitting the τ-values and defining $\tilde{\xi}_j := h_1 - h_j$ with the corresponding time scale $\tau_j - \tau_1$ and $j = 2, \ldots, N$, this equation simplifies to

$$p(h^*|h_1, \tau_1; ...; h_N, \tau_N) = \frac{p(h^*)}{p(h_1)} \times \frac{\prod_{i=1}^{N-1} p(\xi_i | \xi_{i+1}; h^*)}{\prod_{i=2}^{N-1} p(\tilde{\xi}_i, |\tilde{\xi}_{i+1}; h1)} \times \frac{p(\xi_N | h^*)}{p(\tilde{\xi}_N | h_1)}. \quad (12)$$

For a given height h^* the probability of its occurrence $p(h^*|h_1, \tau_1; ...; h_N, \tau_N)$ is given by the simple conditional PDFs, which can be calculated from the Fokker-Planck equation, or which can be estimated directly from the data. Note the simple conditional PDFs $p(\xi_i, \tau_i | \xi_j, \tau_j; h^*)$ only contain information about three height values h^*, h_i, h_j; or more abstractly, of three points of the time series $h(t)$. Thus Eq. (12) is a three point closure of the multi-point problem.

3. Results based on observational data

The wave measurements used in this study were taken in the Sea of Japan, at a location 3 km off the Yura fishery harbor, where the water depth is about 43 meters, further details can be found in [27, 28, 29, 30]. First we want to examine if the conditional PDFs depend on the wave height itself by comparing both sides of the equation

$$p(\xi_1 | \xi_2; h^*) = p(\xi_1 | \xi_2). \quad (13)$$

In Fig. 1 the comparison of conditional PDFs from both sides of Eq. (13) at scales $\tau_1 = 14$ and $\tau_2 = 28$ seconds and for two different values of h^* are shown. To get sufficient data we always use an interval of h^* with $\pm \sigma_h / 4$ (where $\sigma_h = \sqrt{\langle h^2 \rangle}$). For $h^* = 0$ in Fig. 1(a) both distributions are almost the same but for values of $h_0 \neq 0$, like in Fig. 1(c) a significant shift of the red contour plot (solid lines), which is the left hand side of Eq. (13), is found. As .epsa result from this one can clearly deduce that the conditional PDFs $p(\xi_1 | \xi_2; h^*)$ do depend on h^*.

Next the Markov properties according to Eq. (7) can be checked. Note that we have to compare two data sets according to $\xi_j |_{\xi_{j+1}; h^*}$ and $\xi_j |_{\xi_{j+1}; \xi_{j+2}; h^*}$, and the size of each of these data sets is very different. The verification is thus performed by the use
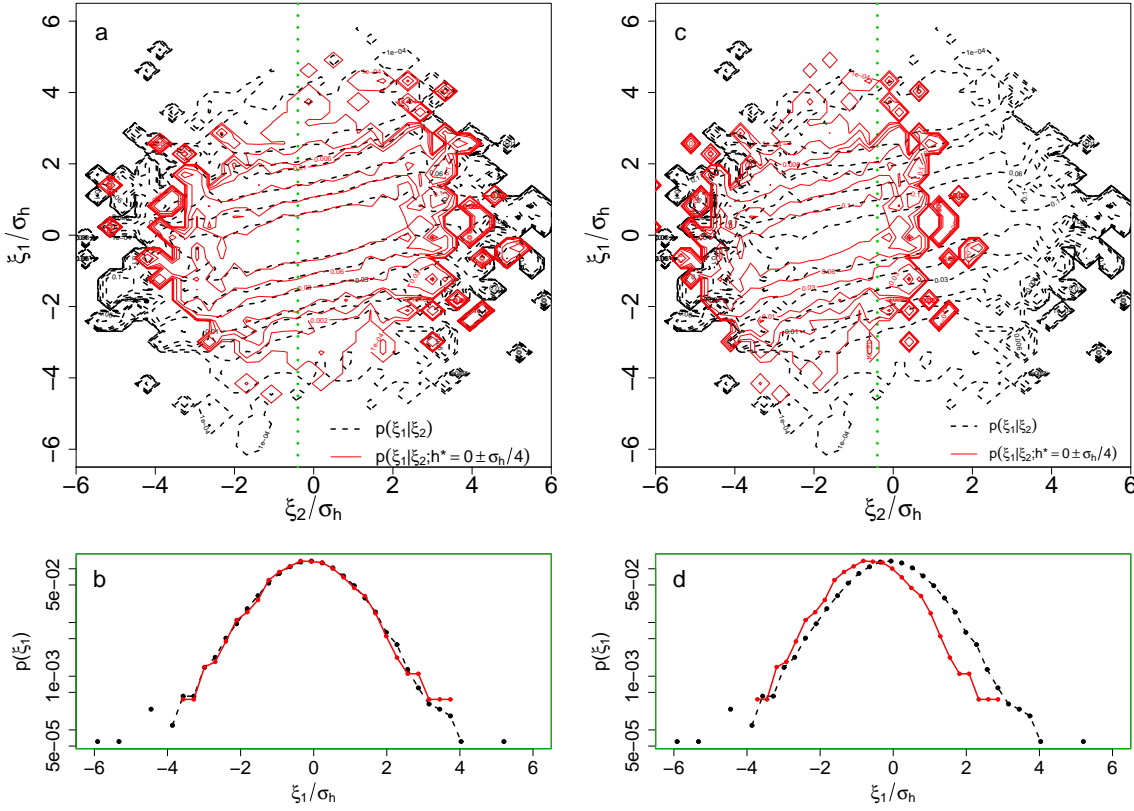
Figure 1. Contour plots of the conditional PDFs $p(\xi_1|\xi_2)$ (dashed black lines) and $p(\xi_1|\xi_2; h^*)$ (solid red lines) with $\xi_j \equiv \xi(\tau_j)$ for scales $\tau_1 = 14$ and $\tau_2 = 2\tau_1$, for $h^* = 0 \pm \sigma_h/4$ (a) and $h^* = 2\sigma_h \pm \sigma_h/4$ (c). Cuts through the conditional PDFs for fixed values of $\xi_2 = -0.5\sigma_h$ in (a) and (c) are shown in (b) and (d) respectively.

of the Wilcoxon test [16] as this test is suitable to compare the statistical similarity of two sample sets of different sizes. The validity of the Wilcoxon test can be shown by the normalized expectation value $< \Delta Q^* >$ of the number of inversions of the conditional wave height increments $\xi_j|_{\xi_{j+1};h^*}$ and $\xi_j|_{\xi_{j+1};\xi_{j+2};h^*}$. If Markov properties are given, $< \Delta Q^* >$ has a value of $\sqrt{2/\pi} \approx 0.8$. The values of $< \Delta Q^* >$ in Fig. 2 for different values of h^* show that Markov properties hold for ($\tau \geq 14$ seconds). This defines a finite minimum step size or scale in the Markov process of the evolution of the surface elevation increments ξ_i. Such a finite step size is well known for stochastic processes in general [31], and the scale is called the Einstein-Markov length, which has for example also been found in a similar way for turbulent flow data, cf. [32, 33]. The scale has been marked by a vertical red dashed line in Fig. 2. Note that compared to our previous work [25] the Markov properties are fulfilled without applying a Hilbert-Huang Transform (HHT) to the original data, which is due to the fact that we have now included the dependencies on h^*.
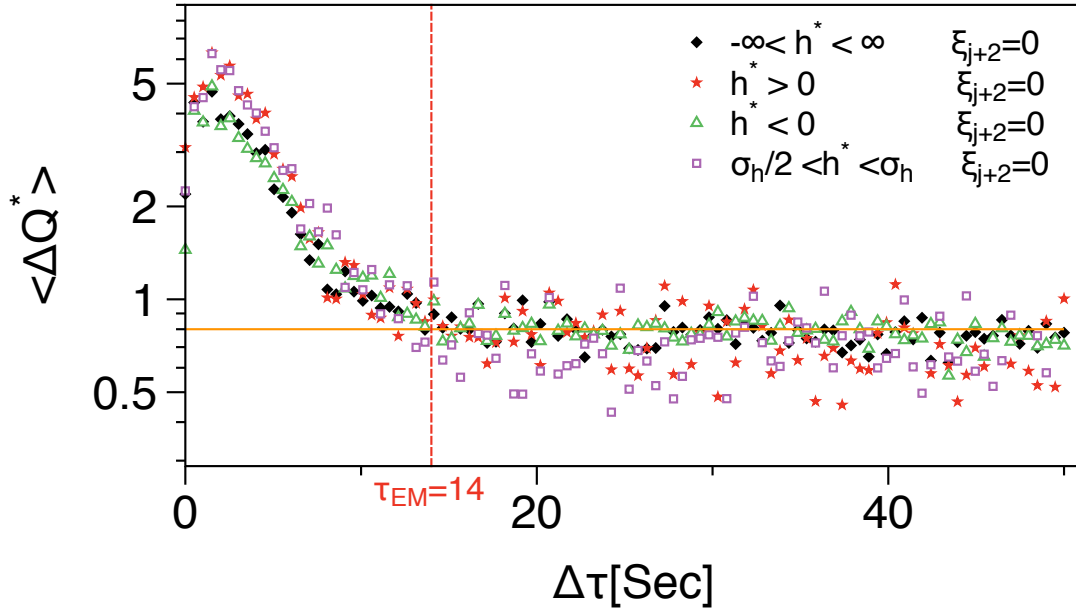
Figure 2. Wilcoxon test of Eq. 7 for different values of h^*.

Based on the finding that Markov properties are fulfilled for the evolution of water surface height increments ξ_j with decreasing time scale τ_j we can now proceed to estimate the corresponding stochastic process via the above mentioned Kramers-Moyal coefficients. Based on the knowledge of the conditional PDF like shown in Fig. 1 the conditional average in Eq. (10) is known too. The estimation of $\lim_{\delta_\tau \to 0}$ causes some problems, in particular due to the Einstein-Markov length, but has been worked out already in several publications [16, 34, 35]. Besides this direct estimation we optimize the obtained functional forms of $D^{(1)}$ and $D^{(2)}$ by minimizing the differences between measured conditional PDFs and those obtained by numerical solutions of the resulting Fokker-Planck equation [36]. Fig. 3 shows the estimated drift and diffusion functions, $D^{(1)}(\xi, \tau, h)$ and $D^{(2)}(\xi, \tau, h)$, of ocean wave surface elevation data for $\tau = 140$ seconds and different values of wave height h^*. For $h^* \neq 0$ the $D^{(1)}(\xi, \tau, h)$ curves are shifted in the vertical direction, whereas no significant change is found for the diffusion term. Furthermore the fourth order Kramers-Moyal coefficient $D^{(4)}(\xi, \tau, h)$ is indeed found to be close to zero for different values of h^*, thus the Fokker-Planck description we propose in Eq. (11) can be assumed to be valid. Note that the surface elevation height increments, ξ, are given in the units of their standard deviation in the limit $\tau \to \infty$, σ_∞, which is identical to $\sqrt{2}\sigma_h \equiv \sqrt{2\langle h^2 \rangle}$ [16].
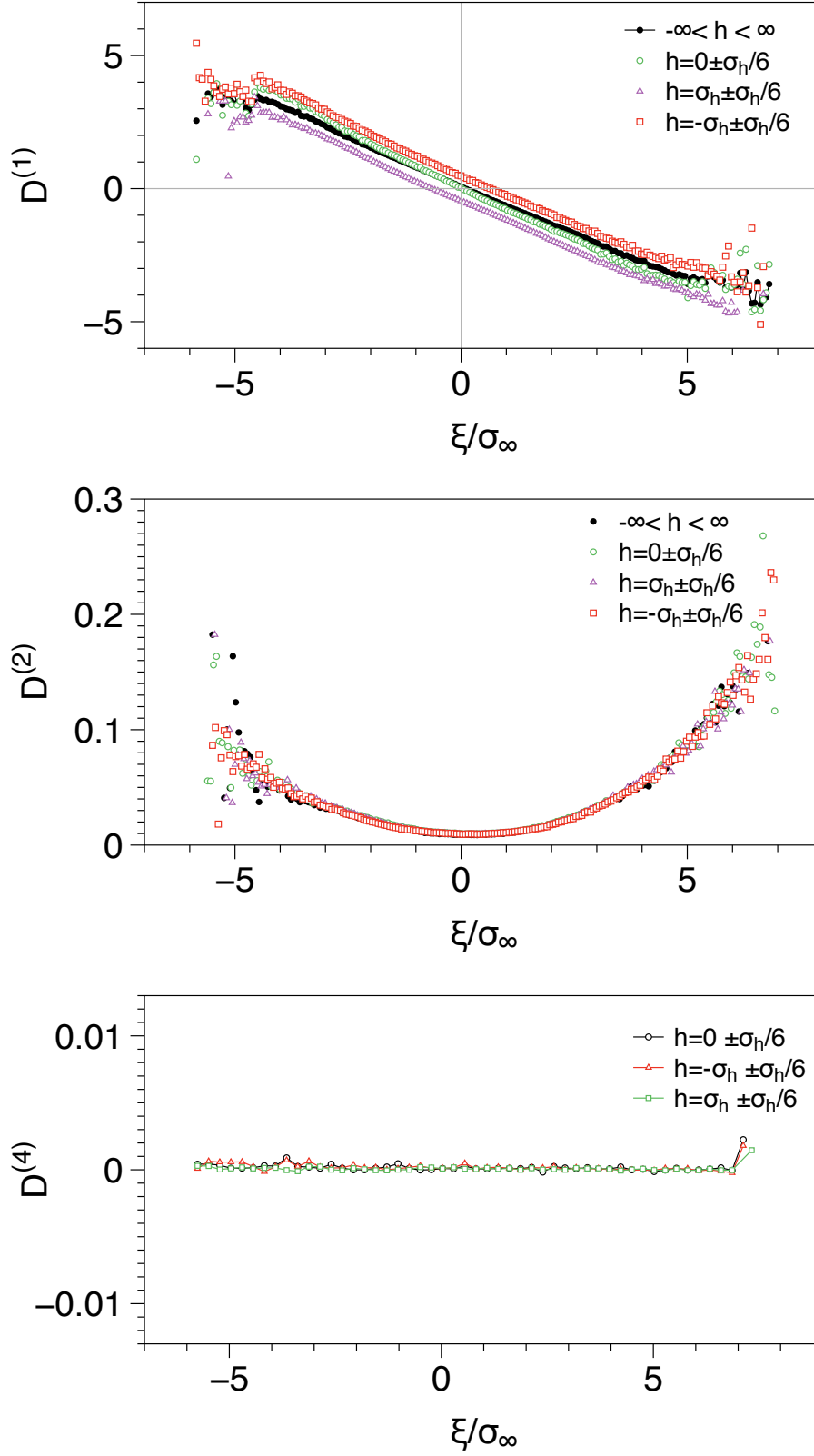
Figure 3. Drift, $D^{(1)}(\xi, \tau, h)$, diffusion, $D^{(2)}(\xi, \tau, h)$, and the Kramers-Moyal coefficient $D^{(4)}(\xi, \tau, h)$ for different value of wave height, h, at $\tau = 140\ s$.

To ease parameterisation, the drift and diffusion terms can be approximated by first and second order polynomials in ξ,

$$D^{(1)}(\xi, \tau, h) = d_{10}(\tau, h) - d_{11}(\tau)\xi,$$
$$D^{(2)}(\xi, \tau) \quad = d_{20}(\tau) - d_{21}(\tau)\xi + d_{21}(\tau)\xi^2. \tag{14}$$

The height dependency of the drift function is expressed by the $d_{10}(h^*, \tau)$-coefficient and our results are shown in Fig. 4. The results indicate once more that we have a strong wave height dependency in our process.
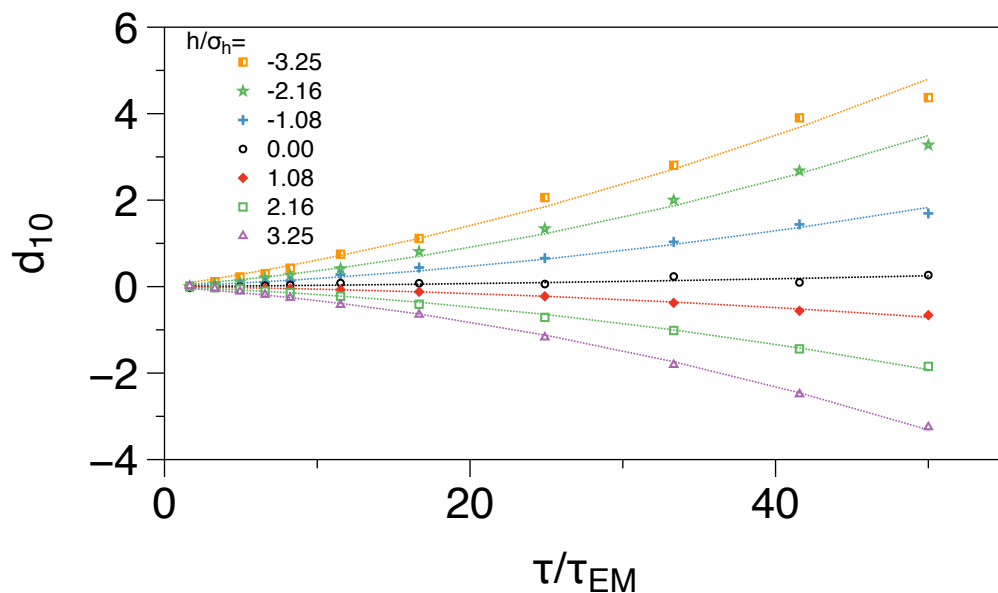

Figure 4. coefficient d_{10} from Eq. 14 as a function of τ for different wave surface height values, h^*. The dotted lines are the second-order polynomial fits in τ.

4. Reconstruction of time series

The knowledge of the conditional probabilities $p(h^*|h_1, \tau_1, \ldots, h_N, \tau_N)$ and its estimation by Eq. (12) can be used to generate a new data point h^*. Shifting the procedure by one step and repeating the same procedure may be used to generate new surrogate time series. For technical reasons one should avoid zeros in conditional pdfs if one uses Eq. (12). Here we used kernel density estimation which is very helpful for parameter ranges for which we have only limited data [37, 38]. The initial idea for reconstructing time-series following this procedure was originally developed in a similar way for fluid turbulence data, see [39]. The time scales we use here for this process are $\tau_n = n \cdot \tau_{EM}$ where $n = 1, 2, \ldots, 7$ and the Einstein-Markov time scale $\tau_{EM} = 14$ seconds, as shown in Fig. 2. (The maximal value of $n = 7$ was chosen, as for that time step the autocorrelation of the height increments approaches zero.) In Fig. 5(a) typical

time series obtained is shown. In the figure part (d) and (e) two selected conditional probabilities $p(h^*|h_1, \tau_1, \ldots, h_N, \tau_N)$ are shown to illustrate our method. In addition to the conditional probabilities the single event probability $p(h^*) = p(h)$ of all height values is shown (red curve). These figures show clearly how the conditional probabilities change with $h_1, \tau_1, \ldots, h_N, \tau_N$ the values of the N wave heights seen before. There are cases when smaller h^*- values are expected in the next step, see Fig. 5 (b), and there are cases when large h^*- values become highly likely, see Fig. 5 (c).
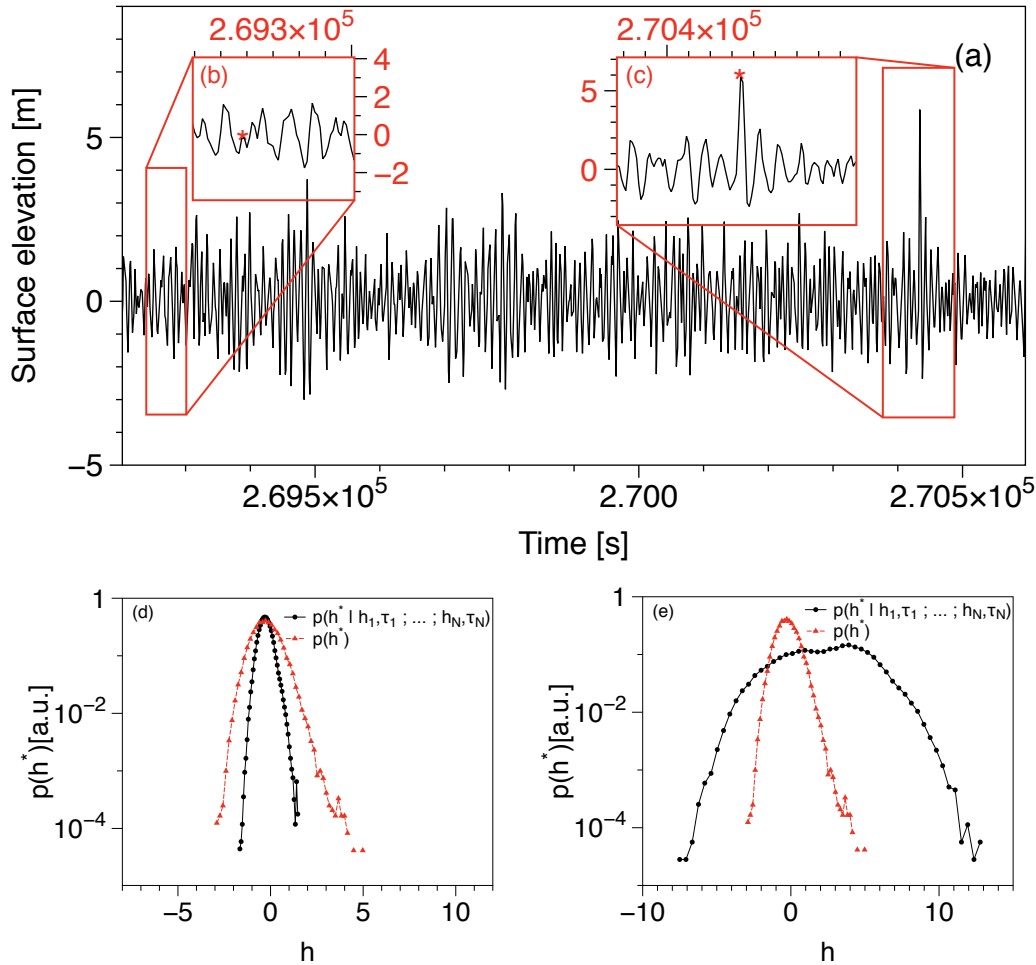

Figure 5. Reconstructed time series (a) after Eq. 12. Two time windows are marked by (*b*) and (*c*) for which the corresponding multi- conditioned PDFs are given (d) and (e). To show the changing volatility the multi- conditioned PDFs (black), the unconditional PDFs (red) estimation from all data are shown too. Note the obvious changes of the likelihood of large wave amplitudes.

To illustrate that the reconstructed time series are indeed statistically similar to the measured wave data we repeat the above mentioned Fokker-Planck analysis. In Fig. 6 we show that from the surrogate data we obtain the same drift and diffusion coefficients. Also the corresponding PDFs $p(\xi_i, \tau_i)$ obtained from the measured data and from the

numerical solution of the Fokker-Planck equation using the estimated drift and diffusion terms are the same as shown in Fig. 7(a). Furthermore the statistics of the wave height maxima are well grasped by the reconstructed data, see Fig. 7(b). Both empirical and reconstructed data follow a generalized gamma distribution very well, as expected from [25]. From this verification of the obtained stochastic process we conclude that both, the empirical data and the reconstructed data, have the same multi-point statistics.
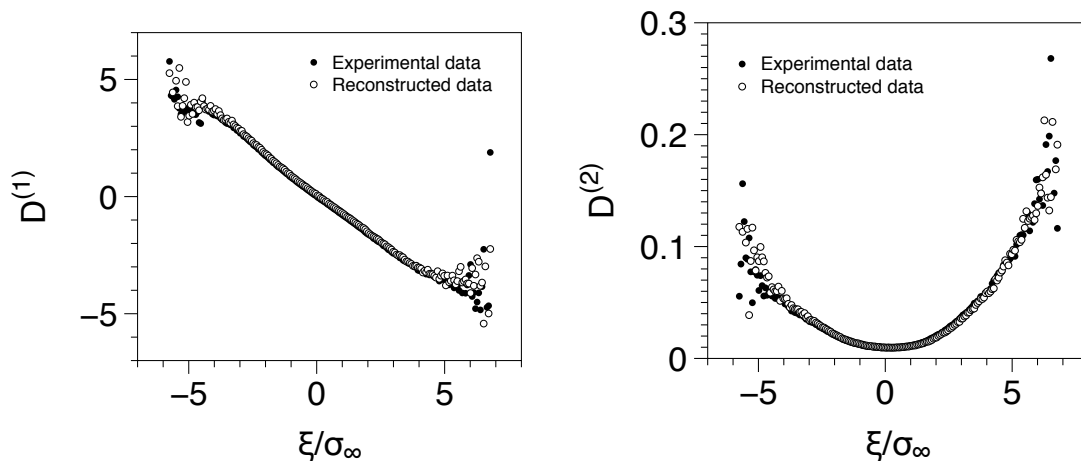

Figure 6. the drift $D^{(1)}(\xi, \tau, h)$ and diffusion $D^{(2)}(\xi, \tau, h)$ coefficients of ocean wave data in time scale $\tau = 280\ s$. The black dots are the original data and hollow circles are from reconstructed time series.

Based on the proposed reconstruction of time series it is now possible to generate long synthetic time series to work out further statistical features of the wave data. We have chosen 1000 data points of empirical data as initial condition and run it to produce 1.1×10^6 synthetic data with sampling rate of $1\ Hz$. In this reconstructed time series we have captured three events that we could consider as rogue waves, using the usual definition [40], saying h^* must be larger than 2 times the significant wave height, which is $2.4\ m$ for our data. The corresponding three sections of the reconstructed time series are shown in Fig. 8 (b,c and d). Also, we performed 4096 different runs of 2048 seconds blocks. From these data we captured 33 time series with extreme values and a corresponding waiting time of about 2.5×10^5 seconds to obtain an extreme event, or a rogue wave.

Next we discuss the possibility to forecast emerging rogue waves. From the conditional probability, Fig. 5(e) (Black curve) we can quantify the likelihood of the appearance of the measured amplitude of $h_r > 5.2\ m$ by integration

$$P_{extreme} = \int_{h_r}^{\infty} p(h^*|h_1, \tau_1; \ldots; h_N, \tau_N)dh^* \tag{15}$$

and obtain 23.6%. This likelihood $P_{extreme}(h_r > 5.2\ m)$ can be evaluated for each time step and result in the changing risk of emerging rogue waves, as shown in Fig. 9. This
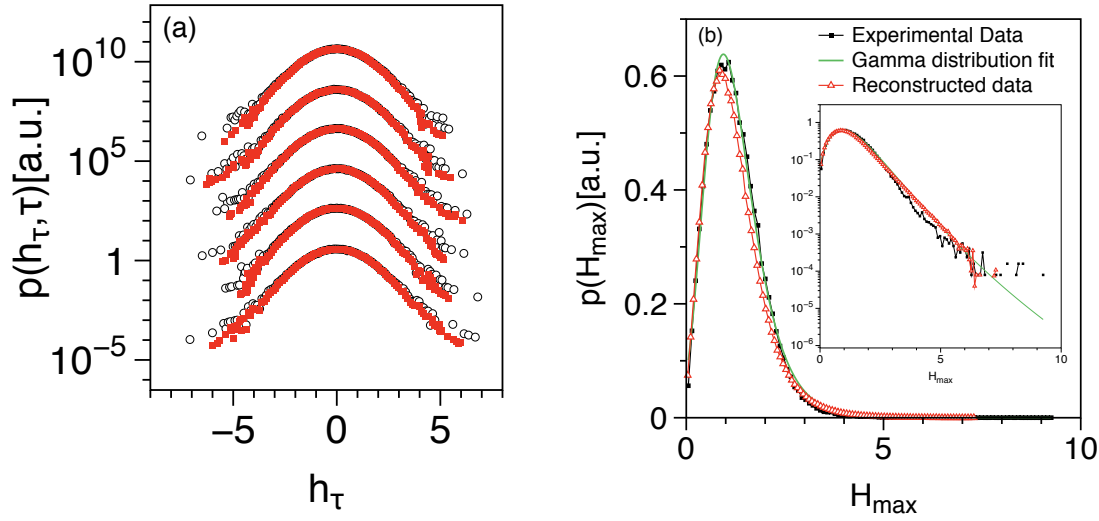
Figure 7. (a) Empirical (hollow circles) and reconstructed (red filled symbols) PDFs for different scales. Time scales $\tau = 14, 28, 42, 56, 70$ and 84 are chosen and PDFs are shifted in vertical direction for clarity of presentation. (b) Distribution of wave height maxima for empirical and reconstructed data in normal and log (inset plot) scale, which follows gamma distribution.
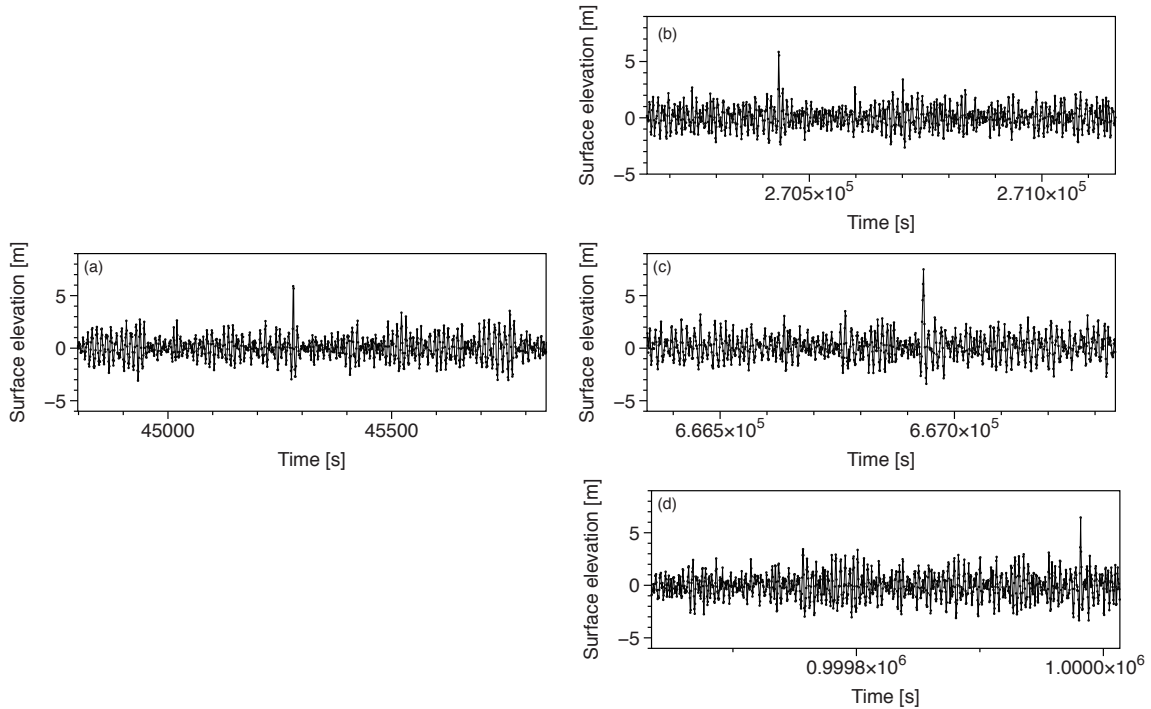

Figure 8. Three different part of stochastic reconstructed time series (b,c and d) based on multi point PDFs, from the empirical data as initial conditions (a).

probabilistic characterization of extreme events returns some false alarms and as well some true hits.
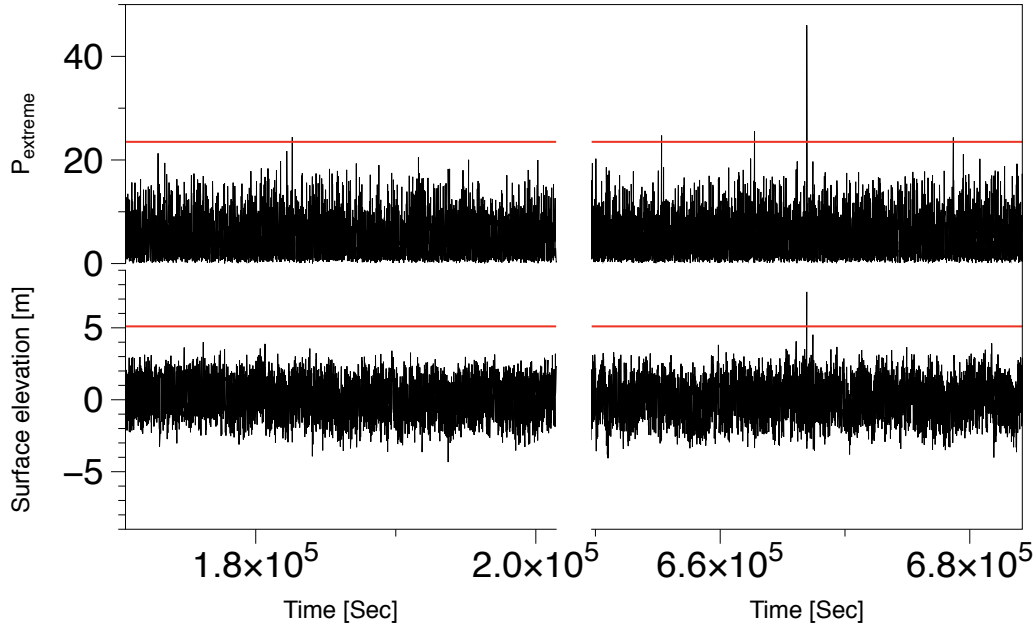
Figure 9. Reconstructed time series (down) and probability of having an extreme event for each reconstructed point, $P_{extreme}$. (up)

A common method to test the quality of a prediction is the receiver operating characteristic curve (ROC) [41, 42, 43]. The idea of ROC consists of comparing the rate of true predicted events with the rate of false alarm. The most quantitative index describing a ROC curve is the area under it which is known as accuracy. In Fig. 10 we have plotted ROC curves for our prediction, $P_{extreme}$, first by considering $h_r = 5.2\ m$ to detect the extreme event alarms. The corresponding ROC curve is plotted in black (solid) line. To investigate the robustness of our reconstruction method, we considered lower amplitude wave height for $h_r = 2.5\ m$ and $h_r = 3.5\ m$ and the corresponding ROC curves are shown in Fig. 10 in red (dotted) and blue (dash) lines, respectively. In all three cases we have accuracy of ROC curves bigger than 80% which indicate that our multi-point procedure is a proper method for time series reconstruction and can be used for short time prediction purposes.
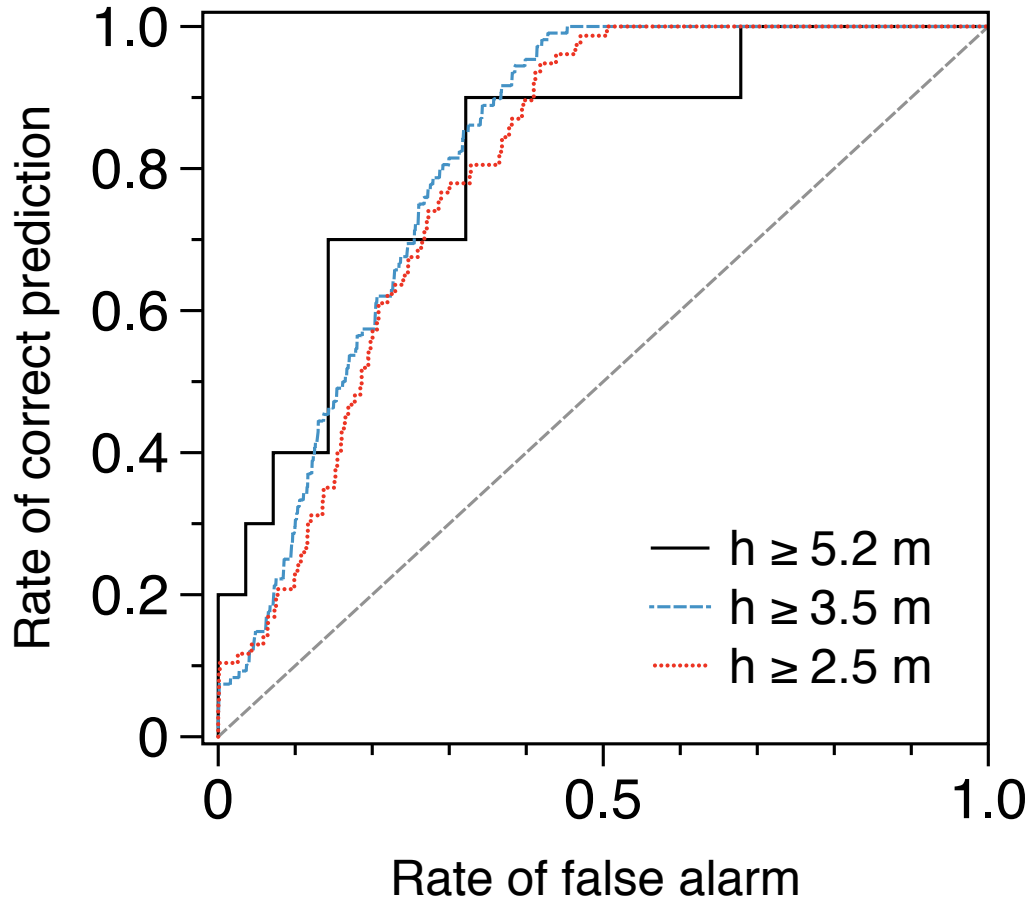
Figure 10. ROC curve for three different estimation of $P_{exterme}$, for $h_r = 5.2\ m$ (solid black line), $h_r = 3.5\ m$ (dash blue line) and $h_r = 2.5\ m$ (dotted red line)

5. Conclusions

We have presented a new approach for a comprehensive analysis of the complexity of ocean wave dynamics. The complexity of multi-point statistics can be simplified by a three point closure, based on which an arbitrary N-Point statistics can be expressed by a hierarchy of nested three point statistics ordered in a cascade like structure. We have been able to show for the first time that by our stochastic approach not only the joint N-point statistics can be grasped but also extreme events, rogue waves, can be captured statistically. We have also shown how for each instant in time the conditional probability of the next wave height can be determined. As the height profile of waves changes from moment to moment, also the probability of the next value of the wave height is changing dynamically. These changes may thus clearly give rise to measures indicating the risk of the appearance of rogue waves ahead of their actual emergence. Most interestingly this was possible although in the measured data only one event of rogue wave was recorded. From our analysis of the occurrence probabilities it becomes

clear that the rogue wave for this wave conditions is integral part of the entire complex stochastic. In our opinion this can only be achieved as we were able to crave out an N-point approach for this complex system.

Acknowledgement

Refrences

[1] T. N. Palmer and J. Raisanen. Quantifying the risk of extreme seasonal precipitation events in a changing climate. *Nature*, 415(6871):512–514, January 2002.

[2] David R. Easterling, Gerald A. Meehl, Camille Parmesan, Stanley A. Changnon, Thomas R. Karl, and Linda O. Mearns. Climate Extremes: Observations, Modeling, and Impacts. *Science*, 289(5487):2068–2074, September 2000.

[3] M. Onorato, S. Residori, U. Bortolozzo, A. Montina, and F.T. Arecchi. Rogue waves and their generating mechanisms in different physical contexts. *Physics Reports*, 528(2):47 – 89, 2013. Rogue waves and their generating mechanisms in different physical contexts.

[4] Ser-Huang Poon, Michael Rockinger, and Jonathan Tawn. Extreme value dependence in financial markets: Diagnostics, models, and financial implications. *Review of Financial Studies*, 17(2):581–610, 2004.

[5] S. El Adlouni, B. Bobe, and T.B.M.J. Ouarda. On the tails of extreme event distributions in hydrology. *Journal of Hydrology*, 355(14):16 – 33, 2008.

[6] V. F. Pisarenko and D. Sornette. Characterization of the frequency of extreme earthquake events by the generalized pareto distribution. *pure and applied geophysics*, 160(12):2343–2364, 2003.

[7] Didier Sornette. Dragon-Kings, Black Swans and the Prediction of Crises, July 2009.

[8] Kung-Sik Chan and Howell Tong. *Chaos: A Statistical Perspective (Springer Series in Statistics)*. Springer, 2001.

[9] N. Akhmediev, Jose M. Soto-Crespo, A. Ankiewicz, and N. Devine. Early detection of rogue waves in a chaotic wave field. *Physics Letters A*, 375(33):2999 – 3001, 2011.

[10] Nail Akhmediev, Adrian Ankiewicz, J.M. Soto-Crespo, and John M. Dudley. Rogue wave early warning through spectral measurements? *Physics Letters A*, 375(3):541 – 544, 2011.

[11] M. Olagnon and M. Prevosto. Are rogue waves beyond conventional predictions? *Proceedings of the 14th Aha Hulikoa Hawaiian Winter Workshop (School of Ocean and Earth Science and Technology (SOEST), Manoa, Hawaii)*, 2005.

[12] Mohammad-Reza Alam. Predictability horizon of oceanic rogue waves. *Geophysical Research Letters*, 41(23):8477–8485, 2014.

[13] A. L. Islas and C. M. Schober. Predicting rogue waves in random oceanic sea states. *Physics of Fluids*, 17(3):–, 2005.

[14] Simon Birkholz, Carsten Brée, Ayhan Demircan, and Günter Steinmeyer. Predictability of rogue events. *Phys. Rev. Lett.*, 114:213901, May 2015.

[15] R. Friedrich and J. Peinke. Description of a turbulent cascade by a fokker-planck equation. *Phys. Rev. Lett.*, 78:863–866, Feb 1997.

[16] C. Renner, J. Peinke, and R. Friedrich. Experimental indications for markov properties of small-scale turbulence. *Journal of Fluid Mechanics*, null:383–409, 4 2001.

[17] C. Renner, J. Peinke, and R. Friedrich. Evidence of markov properties of high frequency exchange rate data. *Physica A: Statistical Mechanics and its Applications*, 298(34):499 – 520, 2001.

[18] A P Nawroth, R Friedrich, and J Peinke. Multi-scale description and prediction of financial time series. *New Journal of Physics*, 12(8):083021, 2010.

[19] Fatemeh Ghasemi, Muhammad Sahimi, J. Peinke, R. Friedrich, G. Reza Jafari, and M. Reza Rahimi Tabar. Markov analysis and kramers-moyal expansion of nonstationary stochastic processes with application to the fluctuations in the oil price. *Phys. Rev. E*, 75:060102, Jun 2007.

[20] R. Friedrich, S. Siegert, J. Peinke, St. Lück, M. Siefert, M. Lindemann, J. Raethjen, G. Deuschl, and G. Pfister. Extracting model equations from experimental data. *Physics Letters A*, 271(3):217 – 222, 2000.

[21] F. Ghasemi, Muhammad Sahimi, J. Peinke, and M.RezaRahimi Tabar. Analysis of non-stationary data for heart-rate fluctuations in terms of drift and diffusion coefficients. *Journal of Biological Physics*, 32(2):117–128, 2006.

[22] F. Atyabi, M. A. Livari, K. Kaviani, and M. R. Tabar. Two statistical methods for resolving healthy individuals and those with congestive heart failure based on extended self-similarity and a recursive method. *J Biol Phys*, 32(6):489–495, Dec 2006.

[23] Rudolf Friedrich, Joachim Peinke, Muhammad Sahimi, and M. Reza Rahimi Tabar. Approaching complexity by stochastic methods: From biological systems to turbulence. *Physics Reports*, 506(5):87 – 162, 2011.

[24] R Stresing and J Peinke. Towards a stochastic multi-point description of turbulence. *New Journal of Physics*, 12(10):103046, 2010.

[25] A Hadjihosseini, J Peinke, and N P Hoffmann. Stochastic analysis of ocean wave states with and without rogue waves. *New Journal of Physics*, 16(5):053037, 2014.

[26] H Risken. *The Fokker-Planck equation : methods of solution and applications*. Springer-Verlag, Berlin New York, 1989.

[27] Nobuhito Mori, Paul C Liu, and Takashi Yasuda. Analysis of freak wave measurements in the sea of japan. *Ocean Engineering*, 29(11):1399–1414, 2002.

[28] Nobuhito Mori and Takashi Yasuda. Effects of high-order nonlinear interactions on unidirectional wave trains. *Ocean engineering*, 29(10):1233–1245, 2002.

[29] Nobuhito Mori and Takashi Yasuda. A weakly non-gaussian model of wave height distribution for random wave train. *Ocean Engineering*, 29(10):1219–1231, 2002.

[30] Nobuhito Mori, Takashi Yasuda, Shin-ichi Nakayama, et al. Statistical properties of freak waves observed in the sea of japan. In *The Tenth International Offshore and Polar Engineering Conference*. International Society of Offshore and Polar Engineers, 2000.

[31] A. Einstein. über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[32] St. Lück, Ch. Renner, J. Peinke, and R. Friedrich. The markoveinstein coherence lengtha new meaning for the taylor length in turbulence. *Physics Letters A*, 359(5):335 – 338, 2006.

[33] R. Friedrich, J. Zeller, and J. Peinke. A note on three-point statistics of velocity increments in turbulence. *EPL (Europhysics Letters)*, 41(2):153, 1998.

[34] Frank Böttcher, Joachim Peinke, David Kleinhans, Rudolf Friedrich, Pedro G. Lind, and Maria Haase. Reconstruction of complex dynamical systems affected by strong measurement noise. *Phys. Rev. Lett.*, 97:090603, Sep 2006.

[35] Julia Gottschall and Joachim Peinke. On the definition and handling of different drift and diffusion estimates. *New Journal of Physics*, 10(8):083034, 2008.

[36] A. P. Nawroth, J. Peinke, D. Kleinhans, and R. Friedrich. Improved estimation of fokker-planck equations through optimization. *Phys. Rev. E*, 76:056102, Nov 2007.

[37] P.J. Green and Bernard. W. Silverman. *Nonparametric Regression and Generalized Linear Models: A roughness penalty approach (Chapman & Hall/CRC Monographs on Statistics & Applied Probability)*. Chapman and Hall/CRC, 1993.

[38] T.J. Hastie and R.J. Tibshirani. *Generalized Additive Models (Chapman & Hall/CRC Monographs on Statistics & Applied Probability)*. Chapman and Hall/CRC, 1990.

[39] A.P. Nawroth and J. Peinke. Multiscale reconstruction of time series. *Physics Letters A*, 360(2):234 – 237, 2006.

[40] N. Akhmediev and E. Pelinovsky. Editorial introductory remarks on discussion & debate: Rogue waves towards a unifying concept?. *The European Physical Journal Special Topics*, 185(1):1–4, 2010.

[41] J. A. Hanley and B. J. McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1):29–36, Apr 1982.

[42] Holger Kantz, EduardoG. Altmann, Sarah Hallerberg, Detlef Holstein, and Anja Riegert. Dynamical interpretation of extreme events: Predictability and predictions. In Sergio Albeverio, Volker Jentsch, and Holger Kantz, editors, *Extreme Events in Nature and Society*, The Frontiers Collection, pages 69–93. Springer Berlin Heidelberg, 2006.

[43] S. Hallerberg and H. Kantz. Influence of the event magnitude on the predictability of an extreme event. *Phys. Rev. E*, 77:011108, Jan 2008.